COMP9444 Neural Networks and Deep Learning

Quiz 6 (Language Processing)

This is an optional quiz to test your understanding of the material from Week 6.

---

1. What are the potential benefits of continuous word representations compared to synonyms or taxonomies?
2. What is meant by the Singular Value Decomposition of a matrix X? What are the special properties of the component matrices? What is the time complexity for computing it?
3. What cost function is used to train the word2vec skip-gram model? (remember to define any symbols you use)
4. Explain why full softmax may not be computationally feasible for word-based language processing tasks.
5. Write the formula for Hierarchical Softmax and explain the meaning of all the symbols.
6. Write the formula for Negative Sampling and explain the meaning of all the symbols.
7. From what probability distribution are the negative examples normally drawn?
8. Explain how an *attention mechanism* can improve the performance of a neural machine translation system.

Make sure you try answering the Questions yourself, before checking the Sample Answers